

文章编号 1004-924X(2026)02-0322-14

改进 U-Net 的全局特征融合水下图像增强网络

高绍姝¹, 焦广森^{1*}, 李广峰¹, 刘宗恩²

- (1. 中国石油大学(华东) 青岛软件学院、计算机科学与技术学院,
山东省智能油气工业软件重点实验室, 山东 青岛 266580;
2. 中国石油大学(华东) 石油工业训练中心, 山东 青岛 266580)

摘要:针对光在水下环境中传播时由于散射和衰减导致水下图像出现颜色偏差和细节模糊问题,提出改进 U-Net 的全局特征融合水下图像增强网络。首先,在编码器和解码器中设计多残差卷积模块对特征信息进行分层次融合处理,减少细节信息丢失。其次,在解码器中引入通道注意力模块对通道进行加权处理,缓解通道退化程度不同的问题。最后,在解码器中设计卷积-置换自注意力模块融合全局信息,促进网络引导图像重建。所提出的方法在 UIEB 数据集上测试,最终在 PSNR, SSIM 和 LPIPS 三个指标上分别取得了 23.42, 0.900 5 和 0.138 5 的成绩,在 LSUI 数据集上测试,最终在 PSNR, SSIM 和 LPIPS 三个指标上分别取得了 29.35, 0.938 2 和 0.088 0 的成绩。实验结果表明所提出的方法在恢复颜色偏差和减少细节模糊方面具有较好的效果,证明其有效性和可行性。

关键词:水下图像增强;深度学习;特征融合;注意力机制;卷积神经网络

中图分类号: TP394.1; TH691.9 **文献标识码:** A

doi: 10.37188/OPE.20263402.0322

CSTR: 32169.14.OPE.20263402.0322

Underwater image enhancement network based on improved U-Net with global feature fusion

GAO Shaoshu¹, JIAO Guangsen^{1*}, LI Guangfeng¹, LIU Zongen²

- (1. Qingdao Institute of Software, College of Computer Science and Technology,
Shandong Key Laboratory of Intelligent Oil & Gas Industrial Software, China University of
Petroleum (East China), Qingdao 266580, China;
2. Petroleum Industry Training Center, China University of Petroleum (East China),
Qingdao 266580, China)

* Corresponding author, E-mail: z23070072@s.upc.edu.cn

Abstract: To address the color deviation and detail blur of underwater images caused by scattering and attenuation when light propagated in the underwater environment, an improved U-Net global feature fusion underwater image enhancement network was proposed. Firstly, a multi-residual convolution module was designed in the encoder and decoder to fuse the feature information hierarchically to reduce the loss of detail information. Secondly, the channel attention module was introduced into the decoder to weight the channels to alleviate the problem of different degrees of channel degradation. Finally, a convolution-per-

收稿日期: 2025-09-27; 修订日期: 2025-11-06.

基金项目: 国家自然科学基金 (No. 61801517)

mutated self-attention module was designed in the decoder to fuse the global information and promote the network-guided image reconstruction. The proposed method was tested on UIEB dataset, and finally achieved 23.42, 0.900 5 and 0.138 5 on PSNR, SSIM and LPIPS, respectively. The results on LSUI dataset were 29.35, 0.938 2 and 0.088 0 on PSNR, SSIM and LPIPS, respectively. Compared with other commonly used underwater image enhancement methods on several public underwater image datasets, the experimental results show that the proposed method has good effect in restoring color deviation and reducing detail blur, which proves its effectiveness and feasibility.

Key words: underwater image enhancement; deep learning; feature fusion; attention mechanism; convolutional neural network

1 引言

清晰的水下图像在各种海洋探索活动中起着至关重要的作用^[1-3],然而,由于光在水下环境中传播时会出现散射和衰减,因此水下图像常受到严重的退化影响,比如颜色偏差和细节模糊^[4-6]。这种退化给水下探索和科学研究带来了巨大的挑战^[7-8],因此,提升水下图像的质量至关重要^[9-10]。

水下图像增强方法通常可以分为两类:传统方法和基于深度学习的方法。传统方法又可以分为两类:无物理模型方法和基于物理模型的方法。无物理模型方法^[11-13]通过修改像素值来提高图像质量。然而,由于这些方法不涉及物理建模过程,因此往往会产生不自然的伪影以及过度或不足的增强效果。基于物理模型的方法^[14-15]试图通过使用一系列先验知识来估计水下成像模型的参数,从而恢复水下图像。然而,由于真实的水下场景复杂多变,单一的增强策略和先验知识无法应对各种不同的水下图像场景。

随着深度学习技术的发展,基于深度学习的图像处理方法成为重要的研究方向^[16-17]。深度学习技术可以弥补传统方法的不足,具有一定的优势^[18-19]。在发展过程中,基于深度学习的水下图像增强方法取得显著进展。Zhou等^[20]提出了一种基于照明区域划分和色彩校正的高性能水下图像增强模型,减少光照不均匀对水下图像质量的影响。Li等^[21]设计了一个无模板的颜色迁移学习框架,通过直接学习迁移参数,解决了传统模板迁移方法需要候选模板的局限性。Ji等^[22]提出了一种由特征提取和特征融合组成的两阶段

网络,设计了一种由嵌套的残差块和多尺度全局信息块组成的高效的级联特征提取结构来提高模型整体性能。Qi等^[23]提出了一种基于Retinex理论的水下图像增强网络,通过融入卷积神经网络和Transformer的组合模块,利用两种架构的优势来改进图像增强效果。Wang等^[24]设计了一种卷积模块和Transformer模块融合的水下图像增强网络,缓解水下环境中的大规模模糊和扩散效应对水下图像的影响。Gao等^[25]提出了一种多级Transformer策略将网络分为高分辨率和低分辨率两个级别,每个分辨率级别分别采用不同的Transformer和注意力计算策略,降低计算资源消耗。

针对水下图像中存在的颜色偏差和细节模糊问题,本文提出改进U-Net的全局特征融合水下图像增强网络。设计多残差卷积模块,使用卷积模块提取不同层次的特征信息,在卷积模块之间使用跳跃连接将不同层次的特征信息进行融合处理,减少特征信息处理过程中的细节信息丢失。引入通道注意力模块,使用注意力机制对通道进行加权处理,根据通道退化程度的不同选择性地增强,缓解水下图像颜色偏差。设计卷积-置换自注意力模块并行处理特征信息,卷积模块提取局部特征信息,置换自注意力模块建立长距离特征依赖关系,两者结果融合处理,获取丰富的全局特征信息引导图像重建,减少水下图像中的细节模糊。

2 本文方法

2.1 总体网络结构

本文提出的水下图像增强网络结构如图1所

示。输入图像首先会经过初始化模块进行处理,得到的结果将会作为编码器模块的输入。在编码器模块中,输入特征会由多残差卷积模块处理,同时空间维度会逐渐减小,通道维度会逐渐增加。通过三层编码器模块来将特征信息映射到潜在空间。跳跃连接模块可以将编码器模块获取的多尺度特征信息传递到对应的解码器模块。解码器模块会将获取到的两个输入在经过拼接后由通道注意力模块处理。接下来,经过处理的特征信息将会被送到卷积-置换自注意力模块中。该模块包含两个子块,两个子块分别对特

征信息进行处理,得到的两个输出在经过元素加法融合后将会被送入下一层。在解码器模块中的每一层都会进行相似的操作。最后,解码器模块的输出在经过输出模块处理后将会得到增强后的图像。

2.2 编码器结构

如图 1 所示,输入图像首先由初始化模块进行处理,该模块包含两个卷积操作模块。由于编码器模块使用了与初始化模块相同的卷积操作模块,因此在接下来介绍编码器模块时会进行详细说明。

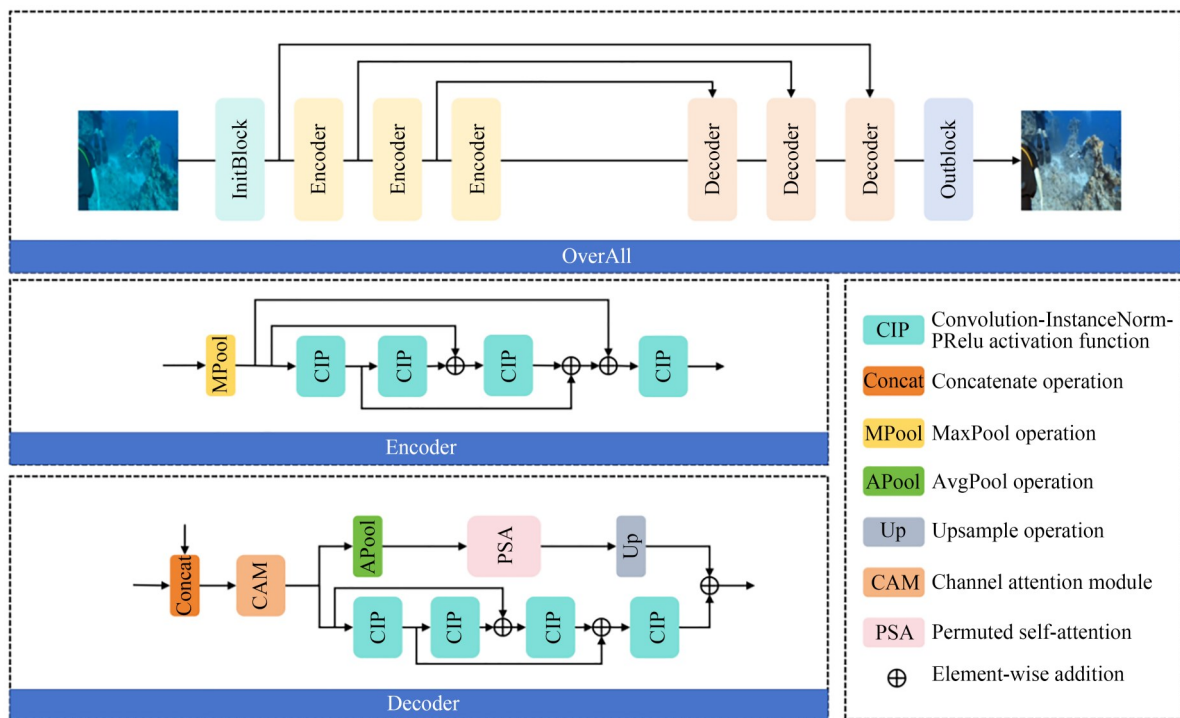


图 1 改进 U-Net 的全局特征融合水下图像增强网络结构

Fig. 1 Underwater image enhancement network structure based on improved U-Net with global feature fusion

图 2 展示了编码器模块的详细结构,编码器模块的每一层都由一个池化层和多残差卷积模块组成。多残差卷积模块包含四个卷积模块,其中在第一层的输入与第二层的输出之间,以及第二层的输入与第三层的输出之间添加了跳跃连接。这种嵌套的跳跃连接策略可以缓解梯度消失的问题并加快模型的学习速度^[18]。之后,使用残差连接来防止网络退化。最后,通过一个卷积模块获得该层的输出。

池化层执行最大池化操作。卷积模块由卷

积层、InstanceNorm 层和 PReLU 激活函数组成。由于水下环境的特殊性, BatchNorm 和 LayerNorm 不适用于水下图像增强任务,而 InstanceNorm 在水下图像增强任务中具有较好的表现^[11],因此在这里使用了 InstanceNorm 来组成卷积模块。

在编码器模块中使用卷积神经网络进行图像特征提取。在此过程中,浅层卷积层主要侧重于获取低级特征信息,包括边缘和简单的纹理信息。相比之下,深层卷积可以提取更高级的语义

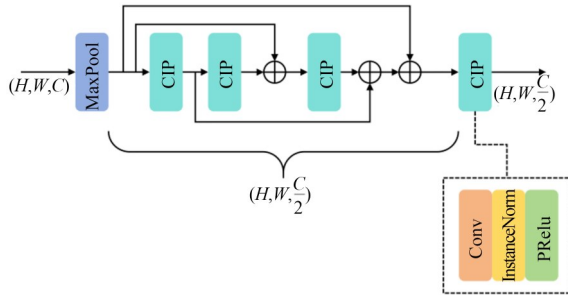


图 2 改进 U-Net 的全局特征融合网络编码器模块网络结构

Fig. 2 Improved U-Net global feature fusion network encoder module network structure

信息。随着特征图的空间维度在每一层逐渐减小,可以提取到更丰富的信息。这种空间尺度的减小是通过下采样实现的,下采样会使原始特征的高度和宽度减半,在一次下采样后空间大小减小为原来的四分之一。

2.3 解码器结构

在水下环境中拍摄的图像往往存在严重的颜色偏差问题,为了缓解这个问题,在解码器模块中引入了通道注意力模块^[26]。具体来说,输入特征首先分别进行平均池化和最大池化操作,得到的结果将分别经过卷积层、ReLU 激活函数和卷积层的处理。两个输出在逐元素相加后,再经过 Sigmoid 激活函数处理,然后将得到的权重与输入特征相乘,得到输出结果。在此选择了平均池化和最大池化相结合的方式而没有仅使用平均池化,是因为最大池化可以突出显著目标特征,这可以使增强后的图像包含更多有用的信息。该过程可以表示为:

$$I_{CA} = I \times S_{Sig} \left\{ M_{CRC} [M_{AP}(I)] + M_{CRC} [M_{MP}(I)] \right\}, \quad (1)$$

其中: M_{AP} 表示平均池化操作, M_{MP} 表示最大池化操作, M_{CRC} 表示经过卷积, ReLU 激活函数和卷积操作, S_{Sig} 表示 Sigmoid 激活函数。

为了应对水下环境中由于大范围的模糊和扩散效应导致的图像质量差的问题,在解码器模块中设计了卷积-置换自注意力模块,以进行全局信息融合并建立长距离特征依赖关系。在解码器模块中,卷积模块和置换自注意力模块并行组合,输入特征分别由卷积模块和置换自注意力模块处理。最后,通过逐元素相加操作将两个子

块的输出进行融合。

传统的自注意力机制在自注意力计算过程中表现出二次复杂度,在处理高分辨率输入图像时对计算和内存资源的需求显著增加。此外,为了减少图像重建过程中图像细节的损失,需要确保每个注意力头生成的注意力图具有代表性。因此,在解码器模块中引入了置换自注意力机制^[27],该机制生成的注意力图具有代表性,同时也在一定程度上降低了自注意力计算复杂度。

如图 3 所示,在进行自注意力计算之前,输入 $X_{in} \in R^{H \times W \times C}$ 被拆分为 N 个方形窗口, S 代表每个窗口的长度,设置为 16。随后,使用三个线性层 L_Q, L_K 和 L_V 来推导出 Q, K 和 V :

$$Q, K, V = L_Q(X), L_K(X), L_V(X). \quad (2)$$

在此过程中, Q 的通道维度保持不变,而 K 和 V 的维度都减小到 $\frac{C}{r^2}$, 得到 $K \in R^{NS^2 \times \frac{C}{r^2}}$ 和

$V \in R^{NS^2 \times \frac{C}{r^2}}$ 。其中, r 表示缩减因子,设置为 2。随后,将 K 和 V 的空间标记排列到通道维度,得到排列后的标记 $K_p \in R^{\frac{NS^2}{r^2} \times C}$ 和 $V_p \in R^{\frac{NS^2}{r^2} \times C}$ 。尽管 K_p 和 V_p 的窗口大小减小到 $\frac{1}{r^2}$,但它们的通道维度仍然保持不变,以此来确保每个注意力头生成的注意力图具有代表性。

置换自注意力模块的公式可以表示如下:

$$M_{PSA}(Q, K_p, V_p) = S_{Softmax} \left(\frac{QK_p^T}{\sqrt{d_k}} + B \right) V_p, \quad (3)$$

其中, B 是可学习的位置嵌入。这种方法使每个标记能够独立进行自注意力计算,从而生成具有代表性的注意力图。对于维度为 $h \times w \times c$ 的特征,传统的窗口自注意力将其分割为 $\frac{h}{w} \times \frac{h}{w}$ 个窗口。每个窗口分别对 Q, K 和 V 进行三次线性投影,并在应用注意力后进行一次额外的线性投影,从而得到 $\Omega(4W^2C^2)$ 。此外,注意力计算需要 $\Omega(2W^4C)$:

$$\Omega(WSA) = 4hwc^2 + 2W^2hwc. \quad (4)$$

置换自注意力模块将 K 和 V 的线性投影和注意力计算的计算量减少到:

$$\Omega(PSA) = 2hw \frac{C^2}{r^2} + 2hwc^2 + 2W^2 \frac{h}{r} \frac{w}{r} C. \quad (5)$$

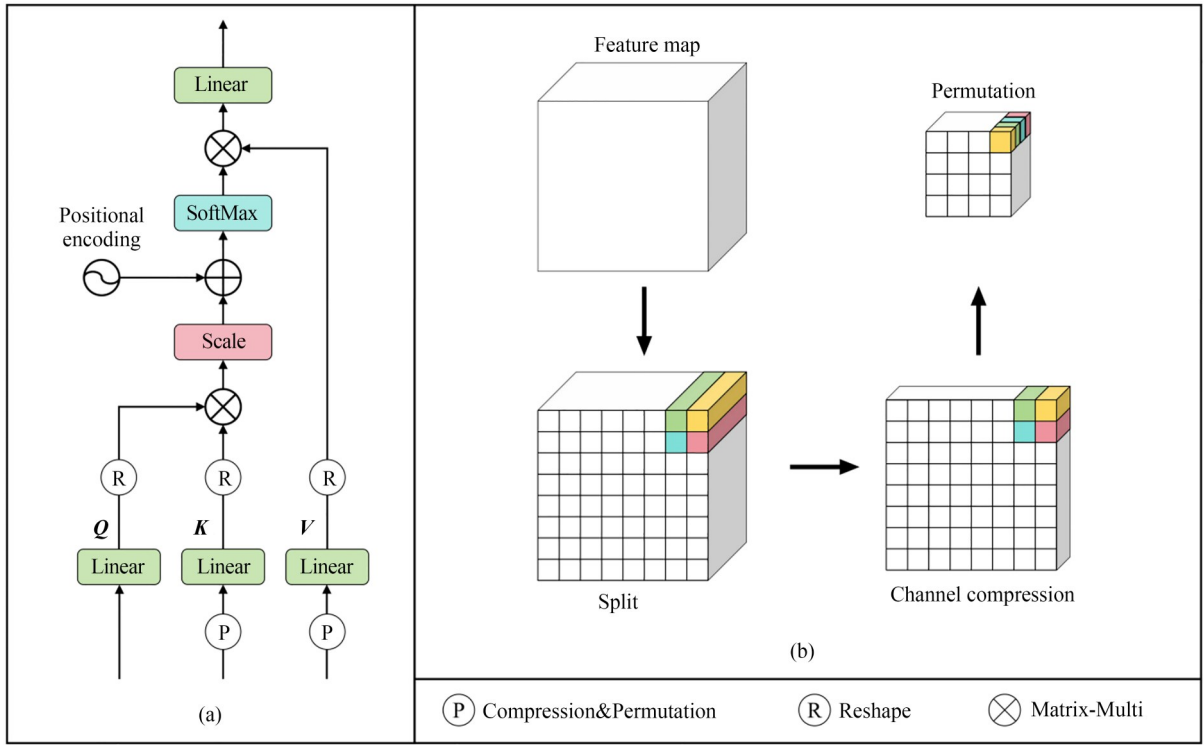


图3 置换自注意力机制

Fig. 3 Permutated self-attention mechanism

置换自注意力模块的详细结构如图4所示。与传统的自注意力机制相比,通过这种方法构建的Transformer模块中使用的自注意力计算方法消耗的计算资源更少。同时,该方法中的每个标记在计算过程中都可以独立参与自注意力计算,从而使得生成的注意力图具有代表性,这在图像重建过程中至关重要。

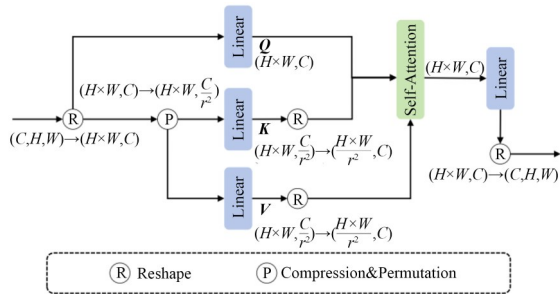


图4 置换自注意力模块网络结构

Fig. 4 Permutated self-attention module network structure

解码器的输出 $X_D \in R^{H \times W \times 64}$ 将作为输出模块的输入进行处理。输出模块包含一个卷积层,该卷积层将 $X_D \in R^{H \times W \times 64}$ 恢复为 $X_{out} \in R^{H \times W \times 3}$,

作为输出图像。

2.4 损失函数

本文网络使用 L_1 损失函数和 MS-SSIM^[28] 损失函数来提高增强图像与参考图像之间的相似度。 L_1 损失函数可表示为:

$$L_1 = \frac{1}{N} \sum_{i=1}^N |X_i - Y_i|. \quad (6)$$

MS-SSIM 是 SSIM^[29] 的改进版本,通过适应图像分辨率和视觉条件的变化,提供了比单尺度 SSIM 更大的灵活性。MS-SSIM 损失函数可表示为:

$$L_M = 1 - L_{MSSIM}(X_i, Y_i). \quad (7)$$

总损失函数可表示为:

$$L_{total} = w_1 L_1 + w_2 L_M, \quad (8)$$

其中, w_1 和 w_2 分别设置为 1 和 0.5。

3 实验结果与分析

3.1 实验环境

本文实验在 Windows 平台进行,使用 NVIDIA RTX A4000 GPU 进行训练。所有输入图像

的大小调整为 256×256 ,使用 AdamW 优化器, batch size 设置为 8, epoch 设置为 300,学习率设置为 0.000 1。

3.2 实验设置

本文实验使用了三个数据集来进行训练和测试。

(1)UIEB^[30]:UIEB 数据集包含 890 张有参考的水下图像和 60 张无参考的水下图像。在有参考水下图像中,70% 的图像作为训练集,20% 作为验证集,10% 作为测试集。测试集包含 89 张图像,命名为 Test-U89。此外,剩余的 60 张无参考水下图像命名为 Test-C60,作为另一个测试集。

(2)LSUI^[31]:LSUI 数据集包含 4 279 张有参考的水下图像,其中选取 70% 的图像作为训练集,20% 作为验证集,剩余的 10% 作为测试集。测试集包含 428 张图像,命名为 Test-L428。

(3)RUIE^[32]:RUIE 数据集包含三个子集,每个子集用于评估不同的增强结果。在此选择了数据集 UCCS 用于评估水下图像增强方法校正颜色偏差的能力。该数据集包含 100 张偏蓝的水下图像、100 张偏绿的水下图像和 100 张蓝绿色

调的水下图像,共计 300 张水下图像。该数据集命名为 Test-UCCS。

为了评估各种方法的性能,本文采用了三种全参考指标:峰值信噪比(PSNR)^[33]、结构相似性指数(SSIM)^[29]和可学习感知图像块相似度(LPIPS)^[34],以及两种无参考指标:水下图像质量评估(UIQM)^[35]和水下彩色图像质量评估(UCIQE)^[36]。对于这些指标,除 LPIPS 外,值越高表示图像质量越好,而 LPIPS 的值越低则表示图像质量越好。

3.3 对比实验

本文将所提出的方法与其他常用方法进行了对比,对比的方法包括 UDCP^[37],SPDF^[38],WaterNet^[30],SCNet^[39],STSC^[40],Semi-UIR^[41],Ushape^[31],CCMSR^[23]和 Phaseformer^[42]。

3.3.1 定量对比

本文在 Test-U89 和 Test-L428 数据集上与其他方法进行了对比,使用三个全参考指标来评估这些方法的有效性,实验结果如表 1 所示。在以上测试集中,本文所提出的方法在 PSNR,SSIM 和 LPIPS 指标上均取得了更好的结果。

表 1 在 Test-U89 和 Test-L428 数据集上的实验结果
Tab. 1 Experimental results on Test-U89 and Test-L428 datasets

Methods	Params/M	FLOPs/G	Runtime/s	Test-U89			Test-L428		
				PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
UDCP	—	—	2.970 0	12.75	0.649 5	0.332 9	14.46	0.734 6	0.340 3
SPDF	—	—	0.780 0	19.76	0.810 2	0.226 7	20.72	0.855 4	0.240 2
WaterNet	1.09	35.71	0.035 2	20.16	0.824 0	0.172 6	22.03	0.875 4	0.216 1
SCNet	0.82	47.11	0.174 0	21.50	0.859 9	0.187 2	21.15	0.876 4	0.216 4
STSC	25.29	373.22	2.697 7	21.87	0.873 8	0.148 4	20.43	0.856 9	0.222 9
Semi-UIR	1.65	72.88	0.110 7	22.56	0.892 4	0.176 1	21.13	0.821 5	0.251 2
Ushape	22.82	5.97	0.147 6	20.89	0.777 5	0.203 9	26.06	0.886 5	0.145 1
CCMSR	21.13	87.20	0.163 4	21.79	0.883 6	0.178 5	20.61	0.842 8	0.240 8
Phaseformer	1.78	26.08	0.165 2	22.40	0.877 3	0.141 6	20.90	0.869 0	0.193 3
Ours	20.68	223.54	0.112 4	23.42	0.900 5	0.138 5	29.35	0.938 2	0.088 0

由于 Test-C60 和 Test-UCCS 数据集不包含真实参考图像,在这两个数据集上进行评估时使用了无参考指标。在这两个数据集上的实验结

果见表 2。可以看出,本文所提出的方法在两个无参考指标 UIQM 和 UCIQE 上也取得了较好的结果。

表 2 在 Test-C60 和 Test-UCCS 数据集上的实验结果

Tab. 2 Experimental results on Test-C60 and Test-UCCS datasets

Methods	Params/M	FLOPs/G	Runtime/s	Test-C60		Test-UCCS	
				UIQM	UCIQE	UIQM	UCIQE
UDCP	—	—	2.970 0	1.644 4	0.545 8	2.238 5	0.545 3
SPDF	—	—	0.780 0	2.496 4	0.570 4	3.012 9	0.502 2
WaterNet	1.09	35.71	0.035 2	2.691 4	0.580 7	3.103 6	0.573 5
SCNet	0.82	47.11	0.174 0	2.554 3	0.562 1	2.865 0	0.542 4
STSC	25.29	373.22	2.697 7	2.704 1	0.568 8	2.912 2	0.536 2
Semi-UIR	1.65	72.88	0.110 7	2.684 9	0.587 6	2.967 6	0.558 5
Ushape	22.82	5.97	0.147 6	2.548 6	0.559 2	2.983 0	0.557 2
CCMSR	21.13	87.20	0.163 4	2.727 3	0.575 2	2.988 2	0.536 2
Phaseformer	1.78	26.08	0.165 2	2.582 4	0.553 9	2.815 7	0.519 6
Ours	20.68	223.54	0.112 4	2.699 6	0.586 9	3.050 7	0.561 9

3.3.2 定性对比

本文在 Test-U89 和 Test-L428 数据集上将所提出的方法与其他方法进行了对比,如图 5 和图 6 所示。从对比结果中可以发现,UDCP 增强图像效果较差,SPDF 增强后的图像颜色偏向灰白,WaterNet 容易引入伪影,SCNet 增强

后的图像亮度较低,STSC 生成的图像容易出现模糊,Ushape 和 Phaseformer 容易出现额外的颜色偏差,SemiUIR 和 CCMSR 在处理水下图像中存在的蓝绿颜色偏差时效果较差。相比之下,本文所提出的方法增强得到的图像效果更好。

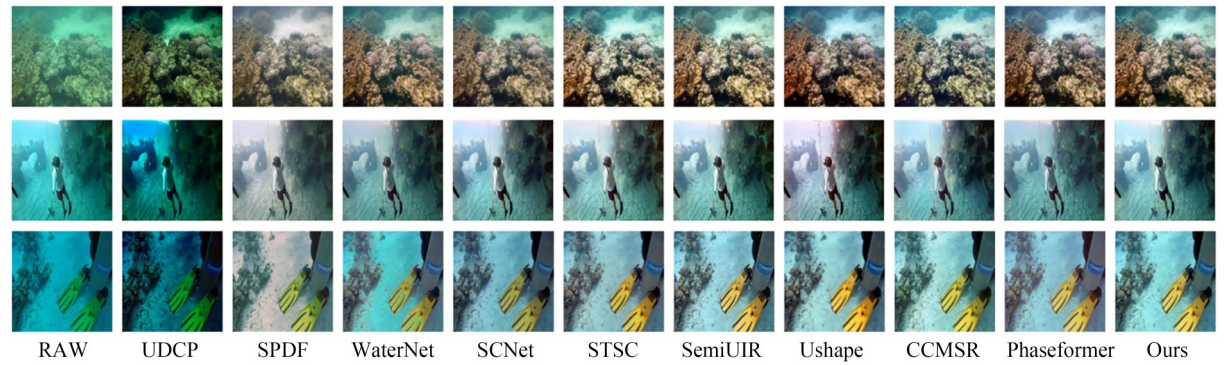


图 5 在 Test-U89 数据集上的对比结果

Fig. 5 Comparison results on the Test-U89 dataset

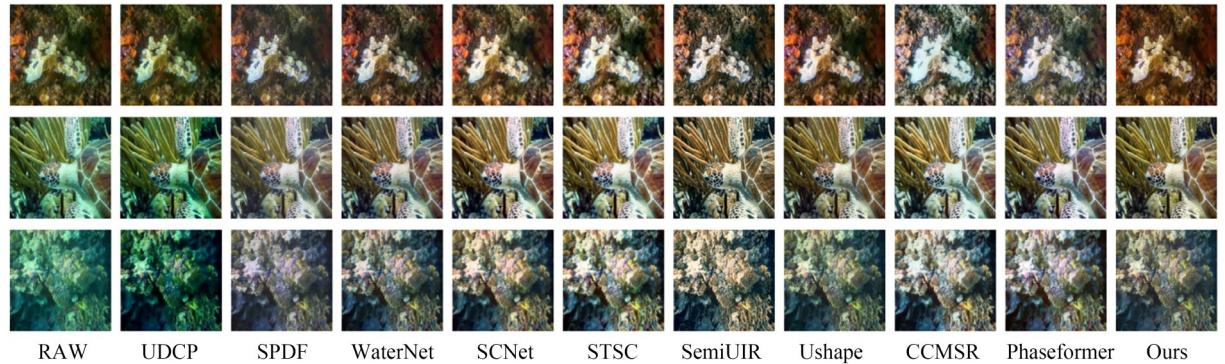


图 6 在 Test-L428 数据集上的对比结果

Fig. 6 Comparison results on the Test-L428 dataset

此外,本文在 Test-C60 和 Test-UCCS 数据集上进行了对比,如图 7 和图 8 所示(彩图见期刊电子版)。这些数据集包含严重退化的图像。在 Test-C60 数据集中分别展示了不同方法对带有黄色、蓝色和绿色色调的图像的增强结果。在 Test-UCCS 数据集中分别展示了不同方法对

带有蓝色、绿色和蓝绿色色调的图像的增强结果。可以看出,本文所提出的方法能够更好地处理水下图像中的各种颜色偏差问题,而其他方法均存在增强效果较差的情况,这证明了本文所提出的方法在恢复颜色偏差方面具有更好的效果。

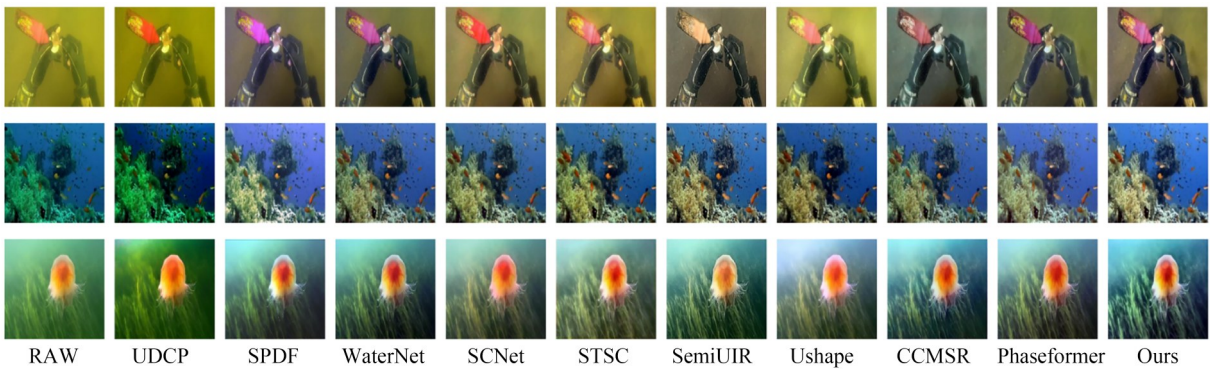


图7 在 Test-C60 数据集上的对比结果
Fig. 7 Comparison results on the Test-C60 dataset

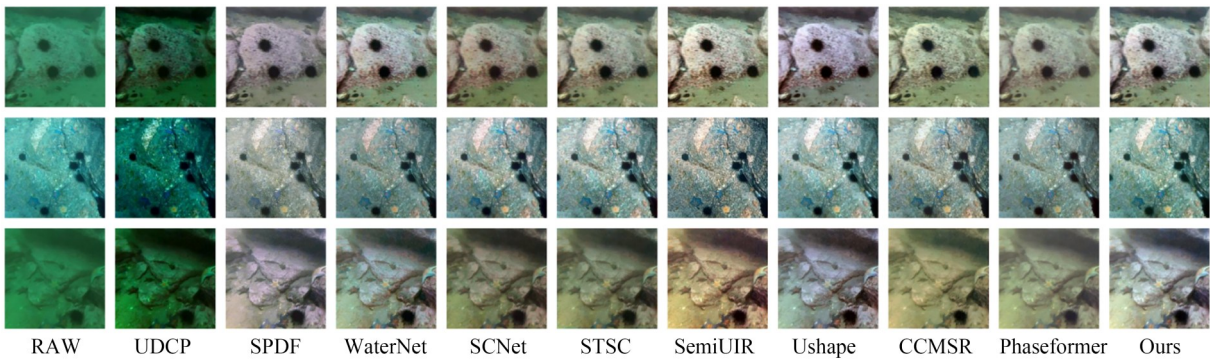


图8 在 Test-UCCS 数据集上的对比结果
Fig. 8 Comparison results on the Test-UCCS dataset

3.4 消融实验

本文在 Test-U89 数据集上进行了消融实验以验证网络中每个模块的作用。

首先对通道注意力模块和置换自注意力模块进行了超参数敏感性实验。

对于通道注意力模块,将其中的降维系数 r 分别设置为 4, 8 和 16, 得到的实验结果如表 3 所示。

对于置换自注意力模块,将其中的窗口大小 w 分别设置为 8 和 16, 将其中的注意力头数量 n 分别设置为 4 和 8, 得到的实验结果如表 4 所示。

表 3 通道注意力模块超参数敏感性实验

Tab. 3 Super parameter sensitivity experiment of channel attention module					
Methods	PSNR	SSIM	LPIPS	UIQM	UCIQE
$r=4$	23.40	0.899 2	0.139 2	3.095 0	0.614 9
$r=8$	23.30	0.902 7	0.140 1	3.122 1	0.615 1
$r=16$	23.42	0.900 5	0.138 5	3.106 7	0.615 4

根据超参数敏感性实验结果,当通道注意力模块中的降维系数 r 设置为 16, 置换自注意力模块中的窗口大小 w 设置为 16, 注意力头数量 n 设置为 4, 此时的网络设置相比于其他的设置在全

表 4 置换自注意力模块超参数敏感性实验

Tab. 4 Super parameter sensitivity experiment of permuted self-attention module

Methods	PSNR	SSIM	LPIPS	UIQM	UCIQE
$w=8, n=8$	23.47	0.899 6	0.139 0	3.094 7	0.613 2
$n=8, n=4$	23.23	0.900 1	0.144 3	3.124 2	0.616 8
$w=16, n=8$	23.37	0.905 4	0.139 3	3.124 1	0.615 6
$w=16, n=4$	23.42	0.900 5	0.138 5	3.106 7	0.615 4

参考指标上取得了最好的效果,同时,在无参考指标上也有较好的效果。因此,最终将网络中通道注意力模块的降维系数 r 设置为16,置换自注意力模块的窗口大小 w 设置为16,注意力头数量 n 设置为4。

之后,为了验证各个模块在网络中的作用,进行了消融实验,对比的网络结构如下所示:

(1) 基础模型(Base)。将编码器模块中卷积模块之间的跳跃连接和残差连接移除,并将卷积模块的数量从四个减少到两个。对于解码器模块中与置换自注意力模块并行的卷积模块也进行了同样的处理。此外,移除了解码器模块中的置换自注意力模块。(2) 基础模型+多残差卷积模块(Base+MRCB)。与基础模型相比,恢复了编码器模块和解码器模块中对卷积模块的所有改进。(3) 基础模型+置换自注意力模块(Base+PSA)。与基础模型相比,恢复了置换自注意力

模块。(4) 完整模型通道注意力模块仅使用平均池化(Full model (AP))。与完整模型相比,通道注意力模块去掉最大池化,仅使用平均池化。

消融实验的结果如表5所示,可以看出,完整模型取得了更好的结果。图9展示了消融实验中不同网络结构的实验结果。未使用多残差卷积模块和置换自注意力模块的基础模型增强图像在细节上效果较差。未使用置换自注意力模块的模型融合全局信息能力较差,增强后的图像出现模糊。未使用多残差卷积模块的模型提取利用有效特征信息的能力较差,增强后的图像出现较多的颜色偏差。仅使用平均池化的通道注意力模块缺少对图像显著信息的提取,增强后的图像无法突出显著目标信息。以上问题在完整模型中得到了缓解,这证明了多残差卷积模块、置换自注意力模块和通道注意力模块设置的有效性和合理性。

表 5 消融实验

Tab. 5 Ablation experiment

Methods	PSNR	SSIM	LPIPS	UIQM	UCIQE
Base	23.10	0.895 6	0.146 4	3.105 9	0.610 5
Base+MRCB	23.18	0.899 6	0.142 4	3.112 0	0.614 5
Base+PSA	23.24	0.898 5	0.144 5	3.104 9	0.613 9
Full model (AP)	23.32	0.900 1	0.140 0	3.108 2	0.612 5
Full model	23.42	0.900 5	0.138 5	3.106 7	0.615 4

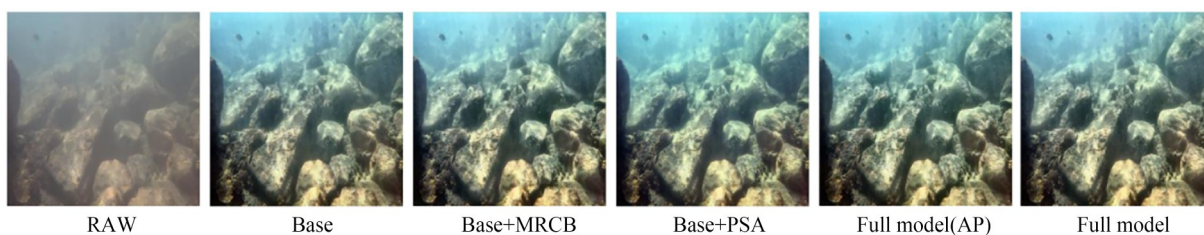


图 9 消融实验中不同网络结构的实验结果

Fig. 9 Experimental results of different network structures in ablation experiment

3.5 应用测试

为评估所提出方法在相关应用中的效果,本文使用 Canny 边缘检测方法、ORB (Oriented FAST and Rotated BRIEF) 特征检测方法和 OTSU 阈值分割方法进行实验。

3.5.1 Canny 边缘检测

本文使用 Canny 边缘检测方法分别对原始图像和增强图像进行处理,实验结果如图 10 所示,图中从左到右分别表示原始图像、原始图像的边缘检测结果、增强图像和增强图像的边缘检测结果。从图中可以看出,增强图像具有更多边缘信息,这证明本文所提出方法可以有效恢复图像的纹理细节信息。

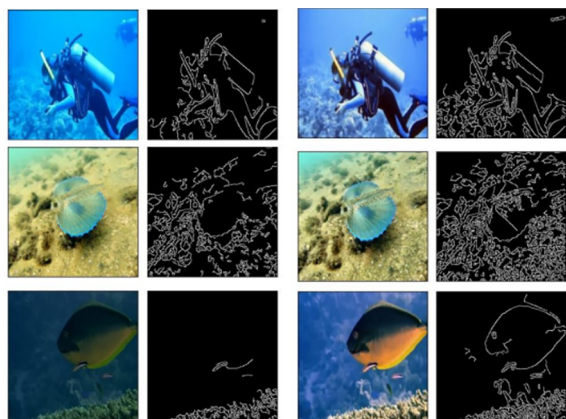


图 10 边缘检测

Fig. 10 Edge detection

3.5.2 ORB 特征检测

本文使用 ORB 特征检测方法对原始图像和增强图像进行测试。首先检测原始图像和增强图像的关键点,之后,将原始图像和增强图像分别与其旋转 180° 的图像进行特征匹配,实验结果如图 11 所示。在图 11 中,左侧表示原始图像与其旋转 180° 的图像进行特征匹配的结果,右侧表示增强图像与其旋转 180° 的图像进行特征匹配的结果,从上到下,原始图像的关键点个数分别为 797, 730 和 838 个,增强图像的关键点个数分别为 1 281, 1 218 和 1 137 个。可以看出,增强图像在保留的特征点个数和匹配对数上均高于原始图像。这表明本文所提出的方法可以有效增强图像。

3.5.3 OTSU 阈值分割

本文使用 OTSU 阈值分割方法分别对原始

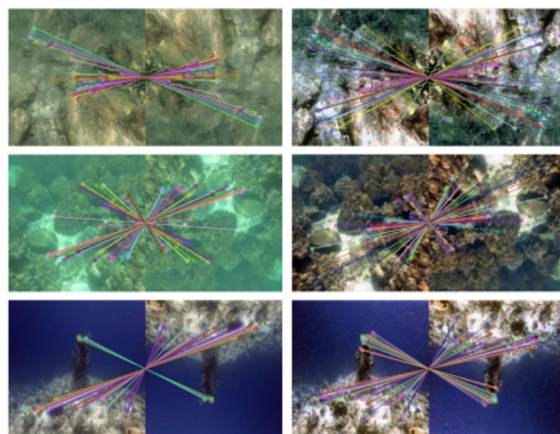


图 11 特征检测

Fig. 11 Feature detection

图像和增强图像进行处理,实验结果如图 12 所示,图中从左到右分别表示原始图像、原始图像经过阈值分割后的结果、增强图像和增强图像经过阈值分割后的结果。从图中可以看出,原始图像由于大范围的模糊,导致分割结果存在轮廓不清晰和内容缺失等问题。增强图像中减少了细节模糊,使得物体与背景更容易区分,分割结果具有更完整的轮廓信息。这证明本文所提出方法在减少背景模糊,提高目标分割精确度上具有良好的效果。

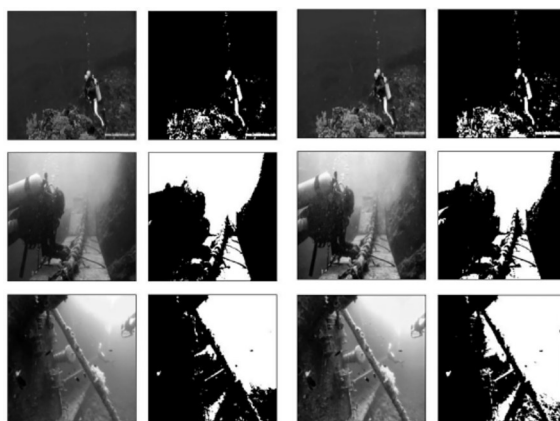


图 12 阈值分割

Fig. 12 Threshold segmentation

4 结 论

本文提出一种改进 U-Net 的全局特征融合水下图像增强网络。在编码器和解码器中设计多残差卷积模块促进特征信息融合,减少在

特征信息处理过程中的细节信息丢失。在解码器中引入通道注意力模块对不同通道进行选择性地增强,减少图像颜色偏差。在解码器中设计卷积-置换自注意力模块实现局部特征信息提取和全局特征信息融合,获取丰富的特征信息引导图像重建,减少图像细节模糊。在UIEB和LSUI数据集上对所提出的方法进行测试,最终在UIEB数据集上PSNR,SSIM和LPIPS三个指标分别取得了23.42,0.900 5和0.138 5的成绩,在LSUI数据集上PSNR,

SSIM和LPIPS三个指标分别取得了29.35,0.938 2和0.088 0的成绩。本文通过实验证明所提出的方法可以有效恢复颜色偏差并减少细节模糊。

作者贡献声明:

高绍姝:写作指导,获取资助,提供资源;
焦广森:提出方法,初稿写作,实验验证;
李广峰:数据分析;
刘宗恩:调查研究。

参考文献:

- [1] 商家硕,李颖,袁靖懿,等. 均值强度差先验驱动的水下图像复原[J]. 光学精密工程, 2025, 33(18): 2962-2979.
SHANG J S, LI Y, YUAN J Y, *et al.* Underwater image restoration driven by the mean intensity difference prior [J]. *Opt. Precision Eng.*, 2025, 33(18): 2962-2979. (in Chinese)
- [2] 李文彪,陶洋,董源,等. 基于颜色先验引导和注意力机制的水下图像增强[J]. 液晶与显示, 2025, 40(8): 1163-1176.
LI W B, TAO Y, DONG Y, *et al.* Underwater image enhancement based on color prior guidance and attention mechanism [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(8): 1163-1176. (in Chinese)
- [3] 程竹明,李佳轩,黄三傲,等. 基于多分支残差注意力网络的水下图像增强[J]. 光学精密工程, 2025, 33(7): 1141-1151.
CHENG Z M, LI J X, HUANG S A, *et al.* Underwater image enhancement based on multi-branch residual attention network [J]. *Opt. Precision Eng.*, 2025, 33(7): 1141-1151. (in Chinese)
- [4] 李明桂,周焕银,龚利文. 基于轻量级U形网络的颜色空间优化水下图像增强方法[J]. 机器人, 2026, 48(1): 163-173.
LI M G, ZHOU H Y, GONG L W. Underwater image enhancement method with color space optimization based on lightweight U-net [J]. *Robot*, 2026, 48(1): 163-173. (in Chinese)
- [5] 郭伟,王曼婷,曲海成. 基于多尺度感知的多维空间融合水下图像增强[J/OL]. 计算机应用, 1-11.
GUO W, WANG M T, QU H C. Multiscale Perception-Based Multidimensional Spatial Fusion for Underwater Image Enhancement [J/OL]. *Journal of Computer Applications*, 1-11. (in Chinese)
- [6] 郭梨,黄智奇,何润丰,等. 一种基于改进DCGAN的水下图像增强方法[J]. 光学学报, 2026, 46(2): 05.
GUO L, HUANG Z Q, HE R F, *et al.* An underwater image enhancement method based on improved DCGAN [J]. *Acta Optica Sinica*, 2026, 46(2): 05. (in Chinese)
- [7] 袁国铭,刘海军,李晓丽,等. 纹理感知联合颜色直方图特征的水下图像增强[J]. 光学精密工程, 2024, 32(13): 2112-2127.
YUAN G M, LIU H J, LI X L, *et al.* Underwater image enhancement using joint texture perception and color histogram features [J]. *Opt. Precision Eng.*, 2024, 32(13): 2112-2127. (in Chinese)
- [8] 陶洋,龚霖霆,周立群. 基于Retinex的可变注意力低照度水下图像增强[J]. 液晶与显示, 2025, 40(3): 481-492.
TAO Y, GONG J T, ZHOU L Q. Variable attention low illumination underwater image enhancement based on Retinex [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(3): 481-492. (in Chinese)
- [9] 张明华,刘佳艺,石少华,等. 基于全局特征提取和提示学习的水下图像增强[J]. 华中科技大学学报(自然科学版), 2025, 53(3): 31-40.
ZHANG M H, LIU J Y, SHI S H, *et al.* Underwater image enhancement based on global feature extraction and prompt learning [J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2025, 53(3): 31-40.

- ture Science Edition), 2025, 53(3): 31-40. (in Chinese)
- [10] 彭子芊,程志远,胡佳豪. 基于色彩校正与注意力机制的水下图像增强[J]. 激光与光电子学进展, 2025, 62(14): 73-81.
PENG Z Q, CHENG Z Y, HU J H. Underwater image enhancement algorithm using color correction and attention mechanism[J]. *Laser & Optoelectronics Progress*, 2025, 62(14): 73-81. (in Chinese)
- [11] ANCUTI C, ANCUTI C O, HABER T, *et al.* Enhancing underwater images and videos by fusion[C]. 2012 *IEEE Conference on Computer Vision and Pattern Recognition*. June 16-21, 2012, Providence, RI, USA. IEEE, 2012: 81-88.
- [12] HOU G J, PAN Z K, WANG G D, *et al.* An efficient nonlocal variational method with application to underwater image restoration[J]. *Neurocomputing*, 2019, 369: 106-121.
- [13] FU X Y, ZHUANG P X, HUANG Y, *et al.* A retinex-based enhancing approach for single underwater image[C]. 2014 *IEEE International Conference on Image Processing (ICIP)*. October 27-30, 2014, Paris, France. IEEE, 2015: 4572-4576.
- [14] FU X Y, FAN Z W, LING M, *et al.* Two-step approach for single underwater image enhancement[C]. 2017 *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. November 6-9, 2017, Xiamen, China. IEEE, 2018: 789-794.
- [15] SONG W, WANG Y, HUANG D M, *et al.* A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration[M]. *Advances in Multimedia Information Processing-PCM 2018*. Cham: Springer International Publishing, 2018: 678-688.
- [16] 赵冬冬,叶逸飞,陈朋,等. 基于残差和注意力网络的声呐图像去噪方法[J]. 光电工程, 2023, 50(6): 230017.
ZHAO D D, YE Y F, CHEN P, *et al.* Sonar image denoising method based on residual and attention network[J]. *Opto-Electronic Engineering*, 2023, 50(6): 230017. (in Chinese)
- [17] 王燕,王宏辉,刘树东,等. 多级特征筛选和任务动态对齐的声呐图像小目标检测[J]. 光电工程, 2024, 51(10): 240196.
WANG Y, WANG H H, LIU S D, *et al.* Small target detection in sonar images with multilevel feature screening and task dynamic alignment[J]. *Opto-Electronic Engineering*, 2024, 51(10): 240196. (in Chinese)
- [18] 陈清江,王炫钧,邵菲. 基于多尺度残差注意力网络的水下图像增强[J]. 应用光学, 2024, 45(1): 89-98.
CHEN Q J, WANG X J, SHAO F. Underwater image enhancement based on multiscale residual attention networks[J]. *Journal of Applied Optics*, 2024, 45(1): 89-98. (in Chinese)
- [19] 贺锋涛,吴倩倩,杨祎,等. 基于深度学习的激光散斑图像识别技术研究[J]. 激光技术, 2024, 48(3): 443-448.
HE F T, WU Q Q, YANG Y, *et al.* Research on laser speckle image recognition technology based on transfer learning[J]. *Laser Technology*, 2024, 48(3): 443-448. (in Chinese)
- [20] ZHOU J C, GAI Q L, ZHANG D H, *et al.* IACC: cross-illumination awareness and color correction for underwater images under mixed natural and artificial lighting[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 4201115.
- [21] LI K Q, FAN H T, QI Q, *et al.* TCTL-net: template-free color transfer learning for self-attention driven underwater image enhancement[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, 34(6): 4682-4697.
- [22] JI X, WANG X, HAO L Y, *et al.* CFENet: Cost-effective underwater image enhancement network via cascaded feature extraction[J]. *Engineering Applications of Artificial Intelligence*, 2024, 133: 108561.
- [23] QI H, ZHOU H Y, DONG J Y, *et al.* Deep color-corrected multiscale retinex network for underwater image enhancement[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 62: 4200613.
- [24] WANG B, XU H Y, JIANG G Y, *et al.* UIE-conformer: underwater image enhancement based on convolution and feature fusion transformer[J]. *IEEE Transactions on Emerging Topics in Compu-*

- tational Intelligence*, 2024, 8(2): 1952-1968.
- [25] GAO Z, YANG J, JIANG F L, *et al.* DDformer: Dimension decomposition transformer with semi-supervised learning for underwater image enhancement[J]. *Knowledge-Based Systems*, 2024, 297: 111977.
- [26] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module[M]. *Computer Vision-ECCV 2018*. Cham: Springer International Publishing, 2018: 3-19.
- [27] ZHOU Y P, LI Z, GUO C L, *et al.* SRFormer: Permuted Self-Attention for Single Image Super-Resolution [C]. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 1-6, 2023, Paris, France. IEEE, 2024: 12734-12745.
- [28] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]. *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003. November 9-12, 2003, Pacific Grove, CA, USA. IEEE, 2004: 1398-1402.
- [29] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [30] LI C Y, GUO C L, REN W Q, *et al.* An underwater image enhancement benchmark dataset and beyond[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4376-4389.
- [31] PENG L T, ZHU C L, BIAN L H. U-shape transformer for underwater image enhancement [J]. *IEEE Transactions on Image Processing*, 2023, 32: 3066-3079.
- [32] LIU R S, FAN X, ZHU M, *et al.* Real-world underwater enhancement: challenges, benchmarks, and solutions under natural light[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(12): 4861-4875.
- [33] KORHONEN J, YOU J Y. Peak signal-to-noise ratio revisited: Is simple beautiful? [C]. 2012 *Fourth International Workshop on Quality of Multimedia Experience*. July 5-7, 2012, Melbourne, VIC, Australia. IEEE, 2012: 37-38.
- [34] ZHANG R, ISOLA P, EFROS A A, *et al.* The unreasonable effectiveness of deep features as a perceptual metric [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 586-595.
- [35] PANETTA K, GAO C, AGAIAN S. Human-visual-system-inspired underwater image quality measures[J]. *IEEE Journal of Oceanic Engineering*, 2016, 41(3): 541-551.
- [36] YANG M, SOWMYA A. An underwater color image quality evaluation metric[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 6062-6071.
- [37] DREWS P L J, NASCIMENTO E R, BOTELHO S S C, *et al.* Underwater depth estimation and image restoration based on single images [J]. *IEEE Computer Graphics and Applications*, 2016, 36(2): 24-35.
- [38] KANG Y Z, JIANG Q P, LI C Y, *et al.* A perception-aware decomposition and fusion framework for underwater image enhancement [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(3): 988-1002.
- [39] FU Z Q, LIN X P, WANG W, *et al.* Underwater image enhancement via learning water type desensitized representations [C]. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. May 23-27, 2022, Singapore, Singapore. IEEE, 2022: 2764-2768.
- [40] WANG D, MA L, LIU R S, *et al.* Semantic-aware texture-structure feature collaboration for underwater image enhancement [C]. 2022 *International Conference on Robotics and Automation (ICRA)*. May 23-27, 2022, Philadelphia, PA, USA. IEEE, 2022: 4592-4598.
- [41] HUANG S R, WANG K Y, LIU H, *et al.* Contrastive semi-supervised learning for underwater image restoration via reliable bank[C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 18145-18155.
- [42] KHAN M R, NEGI A, KULKARNI A, *et al.* Phaseformer: phase-based attention mechanism for underwater image restoration and beyond [C].

2025 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. February 26

- March 6, 2025, Tucson, AZ, USA. IEEE, 2025: 9618-9629.

作者简介:



高绍姝(1983—),女,山西太原人,博士,讲师,2013年于北京理工大学获得工学博士学位,目前就职于中国石油大学(华东),主要从事图像处理、大数据分析等方面的研究。

E-mail: gaoshaoshu@upc.edu.cn

通讯作者:



焦广森(2001—),男,山东聊城人,硕士研究生,2023年于山东科技大学获得计算机科学与技术专业学士学位,目前就读于中国石油大学(华东)计算机科学与技术学院,主要研究方向为水下图像增强、水下目标检测等。

E-mail: z23070072@s.upc.edu.cn